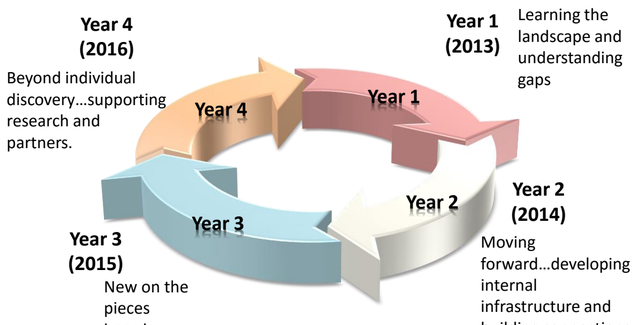
The Ohio State University Libraries

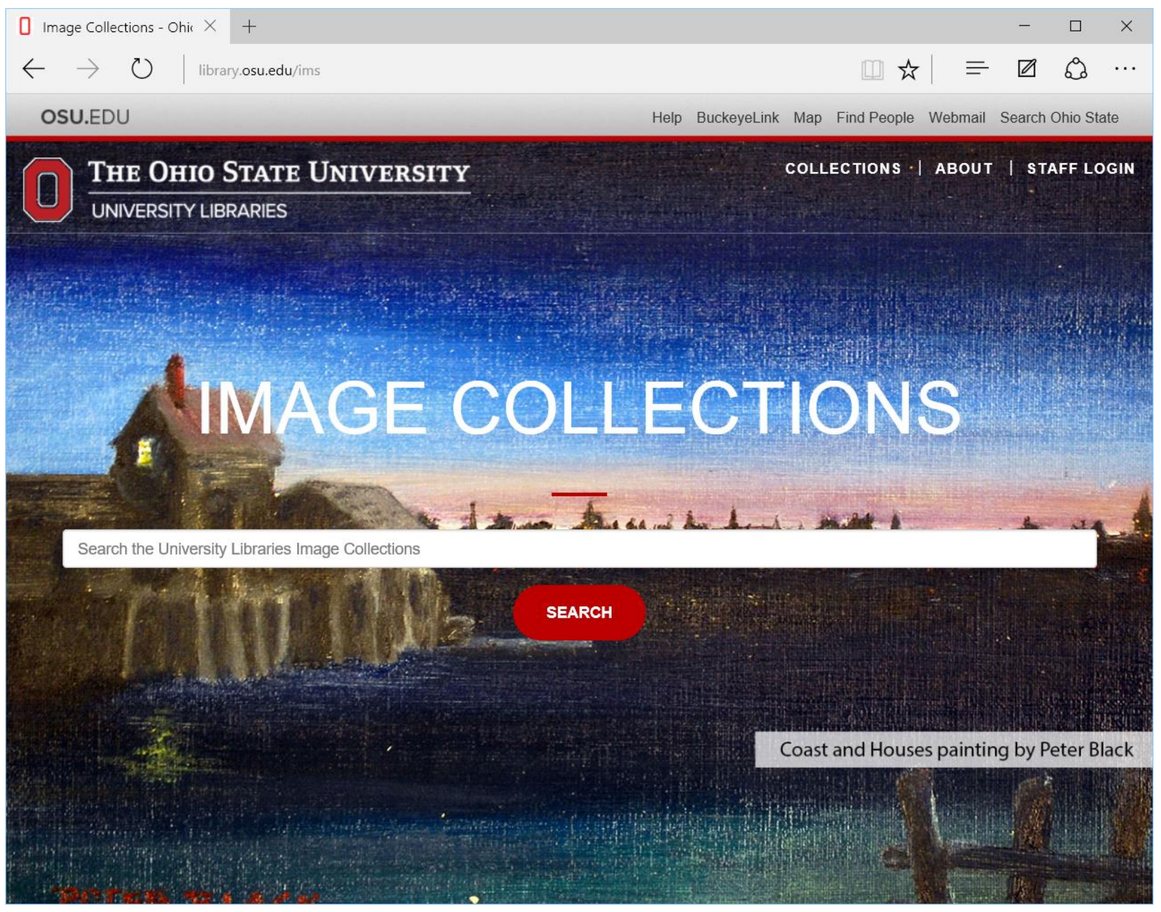
JUMPING HEAD FIRST – OSU'S IMAGE COLLECTIONS MIGRATION

Terry Reese, The Ohio State University Libraries, reese.2179@osu.edu
Ousmane Kebe, The Ohio State University Libraries, kebe.4@osu.edu

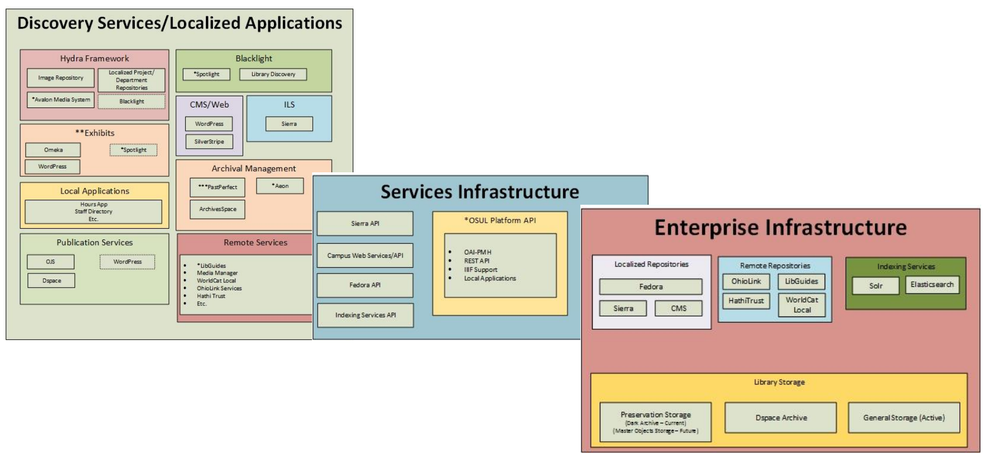
HydraConnect 2015, Minneapolis, Minnesota

DEVELOPING A FOUR YEAR PLAN





MODELLING THE ENVIRONMENT



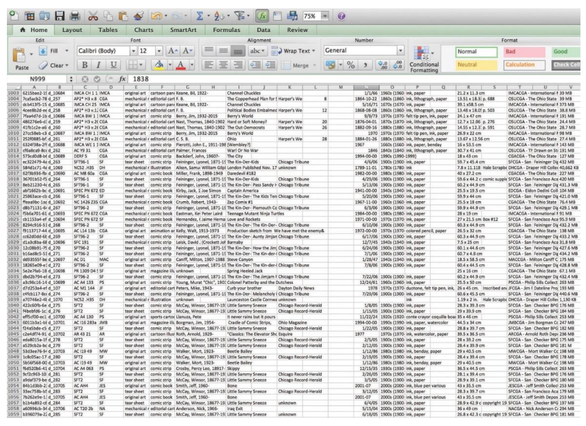
MIGRATING DATA

When we decided to use Fedora/Hydra and Sufia in particular as the framework for Digital Assets Management System, one question which we inevitably had to look into was: How to transfer more than 30, 000 images from our old system to the new one.

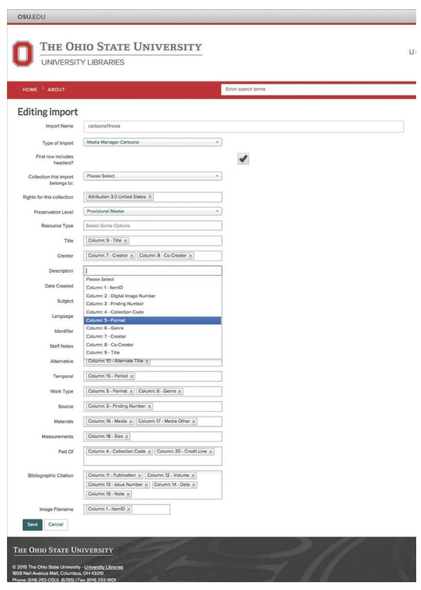
Some of the hurdles we had to overcome included making sure the metadata in our existing sources could be correctly mapped to the metadata fields we defined in our new IMS. We also wanted to ensure that the import process was simple, seamless and accurate. Therefore we wanted a simple easy to use interface along with a step by step workflow that would be easily revertible at any point. Finally we also needed a report with details the success or failure status of each object that went through the import process.

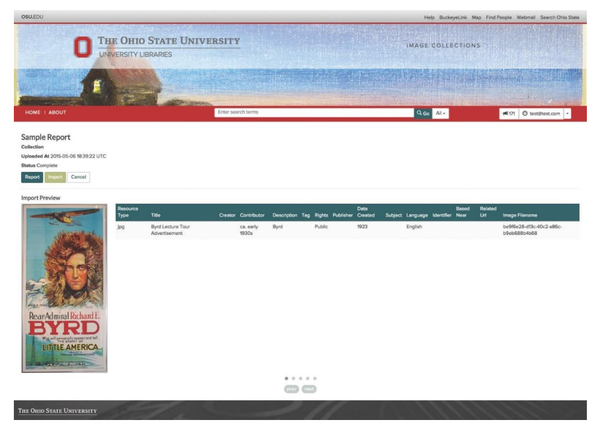
So we embarked in creating a mechanism where we would be able to take a zipped folder of images along with a CSV file which contains the accompanying metadata and ingest the information through Sufia and into Fedora. This process as illustrated below comprises of a few easy steps:

1. Upload CSV file and zipped image folder into system using a provided form.
2. Once files are uploaded you will be presented with a screen that will allow you to map columns in the CSV file to one or more metadata fields in the IMS.
3. Once the mapping process is complete, the CSV file will be read and interpreted and a preview screen will appear giving you the opportunity to browse through your entries before they are ingested to verify that everything looks as expected.
4. After previewing the entries, the ingest process can be started which will read each CSV row and copy all the data into a Sufia GenericFile object according to mapping specified earlier. The accompanying image will also be fetched from the zipped folder and all of it will be ingested into Fedora.
5. After all of the data in the CSV file is read and ingested, you will be able to view a report that will detail the reason for every failed object. Examples of failed object could be images not being found in the zipped folders, or required field not having data in the CSV cell for that specific object.

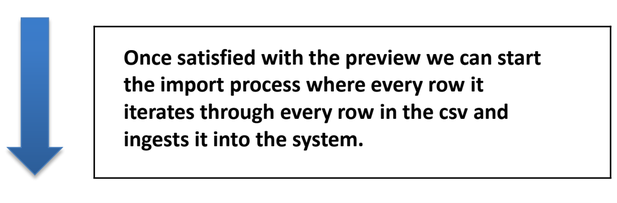


Upload CSV data into system which allows us to map the fields in the csv with fields in our System.

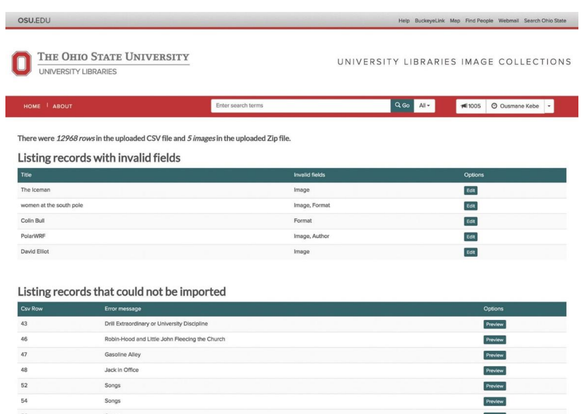


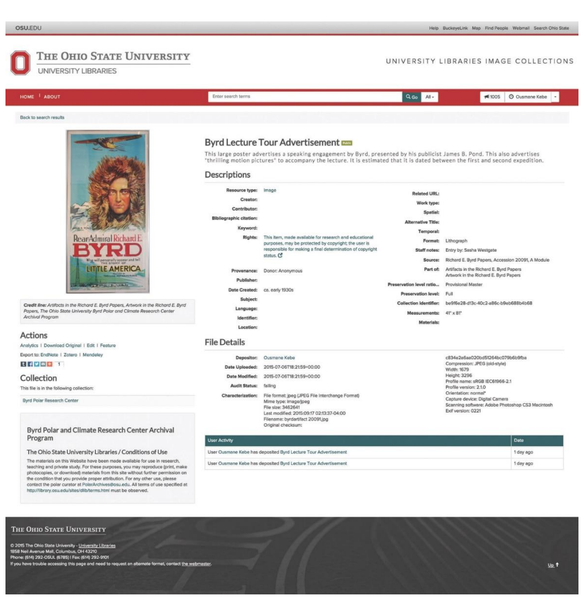


We can map all the fields in the csv file to their appropriate metadata fields and then



Once satisfied with the preview we can start the import process where every row it iterates through every row in the csv and ingests it into the system.





Once import is completed. A report will be available which will detail items which failed the import process along with the reason for the failure.

RESOURCES

- ***Digital Initiatives Principles:*** https://go.osu.edu/osul_di_principles
- ***White Paper:*** https://go.osu.edu/osul_di_whitepaper
- ***Master Objects Repository Report:*** http://go.osu.edu/osul_di_mor
- ***Preservation Framework:*** http://go.osu.edu/OSULDigPrsvPolicy